Clustering and declustering things: The meaning of collective and singulative morphology in Ukrainian*

Marcin Wągiel  
*Masaryk University in Brno & University of Wrocław

Natalia Shlikhutka  
*University of Wrocław

Abstract Many languages have systems of collective and singulative derivational morphology (e.g., de Vries 2021; Dali & Mathieu 2021b). Recent research on Slavic collectives (Grimm & Dočekal 2021; Wągiel 2021a) and singulatives (Kagan & Nurmio to appear; Kagan, Geist & Erschler to appear) shows the significance of these data for the study of linguistic mechanisms of individuation. In this paper, we contribute by investigating the semantics of two derivational morphemes in Ukrainian: the collective suffix \(-j\) and the singulative suffix \(-yn\), and the interaction between the two in secondary singulatives, e.g., pero ‘a feather’ ⇒ pirja ‘(clustered) feathers (as a mass)’ ⇒ pirjina ‘a (single) feather’, and secondary collectives, e.g., popil ‘ash’ ⇒ popelyna ‘a speck of ash’ ⇒ popelynnja ‘(clustered specks of) ash’. Building on the theory of Grimm (2012), we propose a mereotopological account that explains the Ukrainian data in terms of the ontological distinction between integrated objects and clusters: \(-j\) turns properties of integrated objects into properties of clusters, whereas \(-yn\) takes properties of clusters and yields properties of integrated objects.

Keywords: collectives, singulatives, clusters, part-whole structures, mereotopology, Ukrainian

1 Introduction

Collectives are singular nominal expressions that refer to a collection of entities typically conceptualized as a higher-order unit, i.e., a whole with a complex internal structure comprising constituent parts that are themselves independent objects. Due

* We would like to thank the audiences at SALT 33, FDSL 15, DGfS 46, Humboldt University of Berlin and Leipzig University. We are especially grateful to Shweta Alkolkar, Sabine Asmus, Petr Biskup, Ellen Brandner, Alberto Frasson, Berit Gehrke, Ljudmila Geist, Nina Haslinger, Robert Henderson, Olga Kagan, Dorota Klimek-Jankowska, Maria Kouneli, Manfred Krifka, Olav Mueller-Reichau, Liudmyla Petryk, Alexander Pfaff, Paul Pietroski, Božena Rozwadowska, Vesela Simeonova, Ryan Walter Smith and Karolina Zuchewicz for inspiring discussions. All errors are, of course, our own responsibility. We gratefully acknowledge the support of the Centre for Corpus and Experimental Research on Slavic Languages “Slavicus” (SEI) at the University of Wrocław.

©2023 Wągiel, Shlikhutka
Clustering and declustering things
to this dual nature, which involves simultaneously the sense of plurality and singularity, phrases such as (1–2) have received considerable attention in the theoretical literature (e.g., Landman 1989; Barker 1992; Schwarzschild 1996).

(1) a. committee of women
   b. club of graduates
(2) a. pile of dishes
   b. flock of birds

Interestingly, recent work demonstrated the linguistic significance of the distinction between two types of collective nouns: social collectives, which involve an abstract notion of membership, as in (1), and spatial collectives, which describe topological groupings of entities, as in (2) (e.g., Joosten 2010; Pearson 2011; de Vries 2015; Henderson 2017; Zwarts 2020; Wągiel 2021a, 2022; for a recent overview, see de Vries 2021). Especially, the current research on Slavic shows that the operation of spatial collectivization, which we will call CLUSTERING, is encoded in grammar (Grimm & Dočekal 2021; Wągiel 2021a). For instance, the rich inventory of the Czech derivational morphology involves a suffix that forms number-neutral collective nouns that designate topologically structured aggregates of entities denoted by the derivational base, as exemplified in (3).

(3) a. strom ⇒ strom-oví
tree.SG tree-COLL
‘a tree’ ⇒ ‘clump(s) of trees’
b. cihla ⇒ cihl-oví
brick.SG brick-COLL
‘a brick’ ⇒ ‘brickwork’ (Czech; Grimm & Dočekal 2021: 92)

On the other hand, singulatives are derived unit nouns. Typically, they designate a singular object individuated from a plurality perceived as a homogeneous collection of entities, and thus seem to reverse the markedness of the singular/plural distinction, see (4) (e.g., Corbett 2000; Dimmendaal 2000; Grimm 2012, 2018; Mathieu 2012, 2014; Acquaviva 2015; Dali & Mathieu 2021a; Kouneli 2021; for a recent overview, see Dali & Mathieu 2021b).

(4) a. cadair ⇒ cadair-iau
chair.SG chair-PL
‘a chair’ ⇒ ‘chairs’
b. plant ⇒ plent-yn
child.COLL child-SGV
‘children’ ⇒ ‘a child’ (Welsh; Acquaviva 2015: 1174)
What is more important from the perspective of this paper is that in a number of languages singulative morphology can also attach to mass terms in order to derive countable expressions (see, e.g., Mathieu 2012, 2014). In particular, the singulative affix combines with an uncountable noun and triggers a mass-count shift by bringing about a portion reading. For instance, consider (5).

(5)  a. ri.ng’ ⇒ ri.ng’-o
    meat        meat-SGV
    ‘meat’ ⇒ ‘a piece of meat’

   b. kwO.ng’ ⇒ kO.ng’-O
    beer        beer-SGV
    ‘beer’ ⇒ ‘a portion of beer’

In (5a), the singulative suffix -o/O combines with a prototypical mass noun referring to a homogenous body of meat and yields a count noun designating a portion of meat. In a similar vein, in (5b) the singulative morpheme turns an uncountable denotation into a countable one by partitioning it into individuated portions of beer.

Interestingly, recent research on Russian singulatives reveals the significance of the Slavic data in the context of linguistic mechanisms of individuation (Kagan & Nurmio to appear; Kagan et al. to appear). Unlike in languages such as Luo, which allow for singulative morphology to combine with virtually any kind of mass noun, singulative word formation in East Slavic is much more constrained. Specifically, it requires a base with an aggregate meaning in order to apply the individuating operation which we will call DECLUSTERING. This is illustrated in (6).

(6)  a. grad  ⇒ grad-in-a
    hail.SG   hail-SGV-SG
    ‘hail’ ⇒ ‘a hailstone’

   b. soloma ⇒ solom-in-a
    straw.SG   straw-SGV-SG
    ‘straw’ ⇒ ‘a straw’

The examples above show that the Russian singulative suffix -in- selects for an uncountable aggregate noun denoting entities conceptualized as clustered individuals, i.e., spatial configurations of objects that typically either touch each other or remain in close proximity. The attachment of -in- forms a countable noun designating singular units individuated as separate entities.

Though recent research revealed the theoretical relevance of various types of collective and singulative formations, certain aspects of their meaning are still not well understood. In this paper, we will contribute by examining derived collectives
and singulatives in Ukrainian. The data from this language are intriguing since Ukrainian allows for the formation of what we will refer to as SECONDARY COLLECTIVES, i.e., collective nouns formed from derived singulatives, and SECONDARY SINGULATIVES, i.e., singulative nouns formed from derived collectives. We will propose a mereotopological analysis on which the collective morpheme forms a property of clusters of integrated objects by introducing certain constraints on the spatial configuration of a plurality thereof. On the other hand, the singulative morpheme is an atomizer of sorts that selects for an aggregate predicate, i.e., a property of entities prototypically conceptualized as clusters, and turns it into a predicate of discrete singular integrated wholes. We show that the operations can also be applied recursively.

The paper is outlined as follows. In §2, we will discuss the Ukrainian data, specifically we will explore the distributional and semantic properties of the collective suffix -j- and the singulative suffix -yn- as well as the interaction between the two in word formations where they are stacked to form secondary collectives (-yn- + -j-) and secondary singulatives (-j- + -yn-). In §3, we will describe the key mereotopological notions that will allow for distinguishing between singular integrated objects, pluralities thereof and clustered entities. Based on this ontological distinction, in §4 we will propose a mereotopological analysis that will explain the distribution and semantic properties of the Ukrainian collective and singulative morphology. Finally, §5 will conclude the paper.

2 Data

We will investigate two derivational morphemes in Ukrainian: (i) the collective suffix -j- and (ii) the singulative suffix -yn-. They both have a systematically constrained distribution and combine only with a specific subset of nominal expressions that can be characterized as countable unit nouns and uncountable aggregate nouns, respectively. In terms of their contribution, -j- forms uncountable spatial collectives by introducing clustering, whereas -yn- forms countable unit nouns individuated via declustering. Importantly, it is also possible to shift recursively via the alternating application of the two operations since -j- and -yn- can be stacked in both orders, which gives rise to secondary singulatives and collectives.

Below, we will present generalizations based on our dataset, which consists of frequent derived collectives and singulatives extracted from contemporary dictionaries of Ukrainian, corpus data and the own knowledge of the language by the second author who is a native speaker of Ukrainian. All of the lexical items were also

---

1 For the sake of brevity, we will ignore the morphophonological alternations in our examples: $e : i$ and $x : ș$ are standard sound changes in Ukrainian. Note also that -nj-, -sj-, -tj- and -ćj- are allomorphs of -j- and -in- is an allomorph of -yn- in certain morphophonological contexts. All corpus examples...
consulted with other native speakers.

2.1 The collective suffix -j-

Similar to other Slavic languages, Ukrainian has a rich inventory of affixes dedicated to the derivation of various types of collective nouns (see, e.g., Grimm & Dočekal 2021; Wagiel 2021a). One of them is the suffix -j-, which turns unit nouns into uncountable spatial collective nouns. For instance, consider the examples in (7).

(7) a. pero ⇒ pir-j-a
   feather.SG feather-COLL-SG
   ‘a feather’ ⇒ ‘(clustered) feathers (as a mass)’

   b. kamin’ ⇒ kamin-nj-a
   rock.SG rock-COLL-SG
   ‘a rock’ ⇒ ‘(clumped) rocks (as a mass)’ (Ukrainian)

In (7a), while the derivational base denotes a set of singular feathers, the derived collective noun designates clustered feathers conceptualized as a mass of objects that either touch each other or at least remain in close and stable proximity, e.g., a bird’s plumage or tiny fluffy feathers stuffed in a pillow. Likewise, in (7b) the unit noun refers to individual rocks conceptualized as separate objects, whereas the resulting collective describes a mass of clumped rocks, i.e., unindividuated entities organized in a specific spatial configuration.

The suffix -j- always attaches to a concrete countable noun and forms an uncountable collective. As demonstrated in (8), the base can be pluralized and is compatible with cardinal numerals. On the other hand, the derived collective has no plural forms and cannot combine with cardinals.

(8) a. dva per-a
    two feather-PL
    ‘two feathers’

   b. * dva pir-j-a
    two feather-COLL-PL
    Intended: ‘two clusters of feathers’ (Ukrainian)

An example attested in the corpus is provided in (9). Interestingly, in this sentence the collective pirja combines with the singular form vse ‘all, whole’, which triggers universal quantification over whole objects only with uncountable nouns.2

---

2 For universal quantification in the countable domain, the plural forms of vse and the noun are required (for recent discussions of similar data in other languages, see Wagiel 2021b: Ch. 3.5 and Igel 2021).
Clustering and declustering things

(9) Sovenjata, poky ne vyroste vse pirja na krylax i owls.DIM until NEG grow all.SG feather.COLLSG on wings and xvosti, ne vmijut’ tolkom litaty.
tail NEG can properly fly

‘Baby owls cannot fly properly until all the feathers on their wings and tail have grown.’ (Ukrainian; corpus)

Frequent examples of the collective word formation of the type in question include names of plants and plant parts that are often relatively small and occur in clumps or clusters, see (10), small animals, especially insects, that live in big groups such as swarms, packs and flocks, see (11), and a couple of artifacts that usually form (or at least are stored in) an aggregate, see (12).

(10) PLANTS:
   a. lyst ‘a leaf’ ~ lystja ‘foliage’
   b. vita ‘a branch’ ~ vittja ‘(clump(s) of) branches’
   c. volot ‘a panicle branch’ ~ volotija ‘panicle(s)’ (Ukrainian)

(11) SMALL ANIMALS:
   a. muraxa ‘an ant’ ~ murašnja ‘(swarm(s) of) ants’
   b. komaxa ‘an insect’ ~ komašnja ‘(swarm(s) of) insects’
   c. ptaxa ‘a bird’ ~ ptašnja ‘(flock(s) of) birds’ (Ukrainian)

(12) ARTIFACTS:
   a. hančirka ‘a rag’ ~ hančirja ‘(heap(s) of) rags’
   b. toročka ‘a thread’ ~ toroččja ‘fringe (trim)’ (Ukrainian)

To conclude, -j- always attaches to a concrete countable unit noun that has a non-aggregate meaning and designates singular individuated objects. As a result, it forms an uncountable collective noun that refers to clustered entities. The suffix -j- never attaches to mass nouns denoting homogeneous substances and collective nouns.

2.2 The singulative suffix -yn-

Slavic languages, and East Slavic in particular, have rich and productive systems of affixal mass-count shifting (see, e.g., Kagan & Nurmio to appear; Kagan et al. to appear). One such device is the Ukrainian singulative suffix -yn-, which attaches to uncountable aggregate nouns and forms countable unit nouns, as illustrated in (13).
(13)  
a. popil ⇒ popel-yn-a  
    ash.SG  ⇒ ash-SGV-SG  
    ‘ash’ ⇒ ‘a speck of ash’  
b. žyto ⇒ žyt-yn-a  
    rye.SG  ⇒ rye-SGV-SG  
    ‘rye’ ⇒ ‘a grain of rye’  

(Ukrainian)

The suffix -yn- always selects for an uncountable aggregate noun, specifically an expression that denotes a cohesive collection of objects conceptualized as a clustered entity, and forms a countable unit noun that designates a separate object individuated within the relevant collection. This is illustrated in (14), where the aggregate base lacks the plural form and cannot be modified by the cardinal numeral. In contrast, the derived singulative form has the hallmarks of countability, i.e., it can pluralize and felicitously combines directly with cardinals.

(14)  
a. * dva popel-y  
    two ash-PL  
    Intended: ‘two ashes’  
b. dvi popel-yn-y  
    two ash-SGV-PL  
    ‘two specks of ash’  

(Ukrainian)

The sentence in (15) is a corpus example with the diminutive form popelynka. The context makes it clear that the portion of ash in question is conceptualized as a small but clearly individuated object.

(15)  
Popelynka, letivšy z lehotom, sila jomu na ščelepu, vin  
    ash.SGV.DIM.SG was.flying with breeze sat him on jaw he  
    zmaxnuv jiji j posmixnuvsja.  
    blew.away it and he.smiled  
    ‘While flying with the wind, a little speck of ash landed on his jaw; he blew it away and smiled.’  

(Ukrainian; corpus)

Frequent examples of singulatives derived with the suffix -yn- involve names of entities of a similar size and shape that typically occur in clusters. These include small granular objects naturally forming aggregates, see (16), vegetables and fruit that often either grow in bunches or are stockpiled, see (17), precipitation perceived as collections of drops rather than a homogenous body of a substance, see (18), and some artifacts usually stored at one place in large numbers, see (19).

(16)  GRANULAR OBJECTS:
Clustering and declustering things

a. pisok ‘sand’ ∼ piščyna ‘a grain of sand’
b. cukor ‘sugar’ ∼ cukrynka ‘a grain of sugar’
c. perla ‘pearls (as a mass)’ ∼ perlyna ‘a pearl’ (Ukrainian)

(17) VEGETABLES AND FRUIT:
a. kartoplja ‘potatoes (as a mass)’ ∼ kartoplyna ‘a potato’
b. kukurudza ‘corn’ ∼ kukurudzyنا ‘a cob of corn’
c. vynohrad ‘grapes (as a mass)’ ∼ vynohradyna ‘a grape’ (Ukrainian)

(18) PRECIPITATION:
a. snih ‘snow’ ∼ snižyna ‘a snowflake’
b. hrad ‘hail’ ∼ hradyna ‘a hailstone’
c. došč ‘rain’ ∼ doščyna ‘a rain drop’ (Ukrainian)

(19) ARTIFACTS:
a. odeža ‘clothing’ ∼ odežyna ‘an item of clothing’
b. posud ‘dishes (as a mass)’ ∼ posudyna ‘a dish’ (Ukrainian)

To conclude, -yn- always attaches to an uncountable expression. The base is typically an aggregate noun that designates clustered objects and the outcome is a countable unit noun that refers to a singular object individuated within an aggregate. The suffix -yn- does not attach to solid and liquid substance nouns that denote a homogenous body of matter rather than a configuration of separate grains, drops, specks etc.

Having discussed the core semantic properties of the Ukrainian collective suffix -j- and the singulative suffix -yn-, let us now explore the interaction between the two.

2.3 Stacking

A remarkable property of Ukrainian is that it allows for the stacking of the collective and singulative morphology and that it is possible irrespective of the ordering of the morphemes. Admittedly, the cases of such stacking are rare, but we will demonstrate below that both combinations of -j- + -yn- and -yn- + -j- are available in the language. This feature of the Ukrainian system gives rise to complex series involving what we will call secondary collectives, i.e., collectives formed from derived singulatives, and secondary singulatives, i.e., singulatives formed from derived collectives. Let us first discuss the latter.

A series involving a secondary singulative is exemplified in (20). First, the collective suffix -j- attaches to the countable unit noun to derive the spatial collective denoting clustered feathers, recall (7a). Then, the derived collective noun serves as the aggregate base for the formation of the singulative noun by the suffix -yn-.
At first sight, it may seem that the formation of the secondary singulative is redundant since both *pero* and *pirjina* are singular countable nouns referring to feathers. However, native speakers report subtle semantic differences between the two. While *pero* is unmarked and can be used to describe any kind of feather, the meaning of *pirjina* is marked and the noun implies a small (usually fluffy) feather that was most probably recently detached from a bird’s belly or fell out of a pillow.

Just like all other singulatives derived with the suffix -yn-, recall (14), secondary collectives are countable. This is demonstrated in (21).

(21)  
idvі pirjiny  
two feather.COLL.SGV.PL  
‘two (single) feathers’  
(Ukrainian)

An attested example with a pluralized secondary singulative modified by the cardinal numeral is given in (22). The native speaker reading of this sentence emphasizes that the two feathers were detached from a bird’s body.

(22)  
Na podvikonni ležaly dvi pirjiny – čorna ta bila.  
on windowsill lied two feather.COLL.SGV.PL black and white  
‘There were two feathers on the windowsill, a black one and a white one.’  
(Ukrainian; corpus)

Two more examples of complex series showing the formation of secondary singulatives are given in (23).

(23)  
a.  
ptaxa ‘a bird’ ∼ ptašnja ‘(flock(s) of) birds’ ∼ ptašyna ‘a (lone) bird’  
b.  
vолос ‘a hair’ ∼ volossja ‘hair’ ∼ volosyna ‘a (single) hair’ (Ukrainian)

Let us now turn to secondary collectives. An example of this type of word formation is provided in (24). In this case, the order of the collective and singulative morpheme is reversed compared to secondary singulatives. Specifically, -yn- first selects for the uncountable noun with an aggregate meaning and forms the countable unit noun denoting an individuated speck of ash, recall (13a). Subsequently, the collective suffix -j- attaches to the derived singulative noun and forms the collective noun designating an aggregate (or aggregates) of specks of ash, typically clouds of ash.

(24)  
popil ⇒ popel-yn-a ⇒ popel-yn-nj-a  
ash.SG ash-SGV-SG ash-SGV-COLL-SG  
‘ash’ ⇒ ‘a speck of ash’ ⇒ ‘(clustered specks of) ash’  
(Ukrainian)
Again, it might seem that *popil and popelynnja are semantically entirely equivalent. Nonetheless, according to native speakers’ intuitions there is a subtle difference in meaning. While *popil is unspecified and simply refers to ash in any form, popelynnja implies clusters of bigger particles that are easy to individuate, e.g., clumped pieces of burnt wood or clouds of specks of ash in the air.

Similar to other collectives derived with the suffix -j-, recall (8), secondary collectives are uncountable, see (25).

(25) * dva popelynnja
    two ash.COLL.SGV.PL
    Intended: ‘two clusters of specks of ash’ (Ukrainian)

A corpus example is provided in (26), where the secondary collective is modified by an adjective in the singular. The use of popelynnja in the sentence implies clouds of large and thick specks of ash, which enhances the sense of overwhelming evil.

(26) Vin pryjšov z t’manymi nočamy, iz smażnymy vitramy, ščo
    he came with dark nights with burned winds that
    nesly lyxovísne popelynnja.
    carried evil.SG ash.SGV.COLL.SG
    ‘He came with dark nights, with burned winds that carried evil clouds of
    ash.’ (Ukrainian; corpus)

Another example of a series involving a secondary collective is given in (27).

(27) konoplja ‘hemp’ ∼ konoplyna ‘a hemp stalk’ ∼ konoplynnja ‘clustered hemp stalks (as a mass)’ (Ukrainian)

To conclude the empirical part of this paper, the suffixes -j- and -yn- express clustering and declustering, respectively. A remarkable property of Ukrainian is that it allows for stacking of these morphemes in both orders, which results in declustering a clustered entity in secondary singulatives and clustering declustered entities in secondary collectives. Though such word formations might seem redundant, native speakers detect subtle semantic differences that suggest that the traces of the alternating application of the operations in question are preserved in the semantics.

3 Conceptual background

In order to account for the data discussed above, we will adopt a theory of parts and wholes called MEREO-TOPOLOGY, which extends a standardly assumed mereological system with topology (Casati & Varzi 1999; see also Grimm 2012 and Wągiel 2021b). In the assumed framework, the key topological notion is the relation of
CONNECTEDNESS \((C)\), which holds between two individuals. \(C\) is reflexive and symmetric and it is incorporated into the system via so-called bridging principles (Varzi 2007), which ensure fundamental interaction with basic notions and axioms within the mereological component. Most importantly, \(C\) is implied by OVERLAP \((O)\). Hence, while a purely mereological structure is based only on the notion of PARthood \((\sqsubseteq)\), a mereotopological structure involves the interplay between parthood and connectedness.

3.1 Integrated objects

Given the extension of mereology with topological notions, various mereotopological concepts can be introduced. These concepts provide means for modelling subtle distinctions between different spatial configurations of entities. For instance, one can distinguish between integrated objects, scattered entities such as substances and arbitrary sums of objects.

Intuitively, an integrated whole is an individual conceptualized as a solid entity that comes in one piece, i.e., its parts stick together, it has a natural boundary and moves across space along continuous paths (for an overview of the relevance of this category for human cognition, see Wągiel 2021b: Ch. 5). For the purpose of modelling integrated objects in mereotopological terms, the key notion is the property of SELF-CONNECTED \((SC)\) (Casati & Varzi 1999: Ch. 4.2). The definition in (28) states that an individual is self-connected if any two parts that form the whole of that individual are connected to each other.\(^3\)

\[(28)\quad SC(x) = \forall y \forall z [\forall w (O(w,x) \leftrightarrow (O(w,y) \lor O(w,z))) \rightarrow C(y,z)]\]

Although self-connectedness is a key notion, it is insufficient for modelling integrated wholes since it allows also for configurations of externally connected individuals, i.e., entities that merely touch each other. Such configurations can be ruled out by the property of STRONGLY SELF-CONNECTED \((SSC)\) (Casati & Varzi 1999: Ch. 4.3). This property is defined in (29) where INT stands for INTERIOR, which is the sum of internal parts of an entity. In prose, an individual is strongly self-connected if it is self-connected and its interior is also self-connected.

\[(29)\quad SSC(x) = SC(x) \land SC(\text{INT}(x))\]

The two additional notions that should be accommodated in the final definition of an integrated object concern relative integrity and mereological maximality. Importantly, integrity should not be considered in absolute terms but rather evaluated with

\(^3\) For a different attempt to capture the notion of an integrated whole, which is not grounded in mereotopology and even rejects mereology altogether, see Moltmann (1997) and for a critical assessment of this approach, see Pianesi (2002).
Clustering and declustering things

With respect to a particular property. In order to meet these two intuitive requirements, mereotopology allows for developing the property of being \textit{MAXIMALLY STRONGLY SELF-CONNECTED (MSSC)} (Casati & Varzi 1999: Ch. 4.3). The formulation in (30) captures the intuitive category of an integrated whole. In prose, if an entity satisfies \textit{MSSC} relative to a property \(P\), then it is the biggest strongly self-connected entity satisfying \(P\).

\[(30) \quad \text{MSSC}(P)(x) = P(x) \land \text{SSC}(x) \land \forall y[P(y) \land \text{SSC}(y) \land o(y, x) \rightarrow y \subseteq x]\]

With the definition in (30), one can distinguish between integrated objects and other types of individuals. In particular, the mereotopological notion of \textit{MSSC} will enable us to account for the distributional constraints regarding the Ukrainian collectivizing suffix \(-j\)- as well as the semantics of the Ukrainian singulative suffix \(-yn\). Having established the treatment of integrated objects as \textit{MSSC} entities, let us now discuss yet another type of internally structured entity that can be defined within the adopted mereotopological framework.

\subsection*{3.2 Clusters}

Apart from defining integrated wholes, one can also use the connectedness relation \(C\) to derive other types of mereotopologically structured part-whole configurations. For example, it is possible to distinguish between arbitrary sums of entities and clustered individuals, i.e., pluralities involving a topological structure with constituent parts touching each other or remaining in close proximity. Such configurations can be captured via the property of \textit{TRANSITIVELY CONNECTED (TC)} (see Grimm 2012: Ch. 4 for the original notion and Wagiel 2021a for a revised formulation). The definition in (31) characterizes structures in which two individuals are connected via a series of mediating entities without being directly connected to each other. Specifically, entities \(x\) and \(y\) are transitively connected relative to a property \(P\), a connection relation \(C\), and a sequence of entities \(Z\), when all members of \(Z\) satisfy \(P\) and \(x\) and \(y\) are connected through the sequence of mediating entities \(z_i\)s in \(Z\).

\[(31) \quad \text{For a finite sequence } Z = \langle z_1, \ldots, z_n \rangle, \text{ TC}(x, y, P, C, Z) \text{ holds iff } \\
z_1 = x, z_n = y, C(z_i, z_{i+1}) \text{ holds for } 1 \leq i < n \text{ and } P(z_i) \text{ holds for } 1 \leq i \leq n.\]

To illustrate, a collection of entities structured by the \(TC\) relation can be exemplified by a chain of individuals \(a, b\) and \(c\), such that \(a\) is connected to \(b\) but not to \(c\), whereas \(c\) is connected to \(b\) but not to \(a\). Though in such a structure the individuals \(a\) and \(c\) are not directly connected, they are transitively connected via the mediating individual \(b\), which is connected to both \(a\) and \(c\). It is important to emphasize that depending on a particular property, different types of connections may be relevant for construing a cluster. Specifically, one can expect clusters defined in terms of
tangential overlap, i.e., sharing (part(s) of) a boundary, but also configurations of entities simply remaining in stable and predictable proximity. Formally, such configurations can be captured via the notions of EXTERNALLY CONNECTED (EC) and PROXIMATELY CONNECTED (PC), respectively (for the detailed discussion, see Grimm 2012: Ch. 4). For instance, while we perceive the entities in the extension of rice as collections of grains typically touching each other, our conceptualization of the referents of rain is such that the drops merely remain in close proximity without making physical contact. For this reason, we assume that the variation described above is accounted for by the C parameter in the definition in (31).

The TC relation is crucial for our purposes since it enables us to define the notion of CLUSTER (CLSTR) (see Grimm 2012: Ch. 4 for the original definition and Wagiel 2021a for a revised formulation). The formula in (32) characterizes an individual x as a cluster relative to a connection relation C and a property P iff x is a sum of individuals having the property P, which are all transitively connected relative to a subset of Z under the same property P and connection relation C. To illustrate, the plurality a⊔b⊔c in the mereotopological structure discussed above is a cluster since it is a sum of transitively connected individuals.

\[
\text{CLSTR}_C(P)(x) = \exists Z [x = \bigcup Z \land \forall z \forall z' \in Z \exists Y \subseteq Z [\text{TC}(z,z',P,C,Y)]]
\]

The notion of CLSTR introduced in (32) allows for modelling various spatially structured configurations of individuals as complex mereotopological objects. As a result, we can distinguish ontologically between three different kinds of individuals: (i) integrated objects, i.e., wholes conceptualized as singular entities that come in one piece, (ii) arbitrary sums thereof, i.e., pluralities of entities with no topological notions involved, and (iii) clusters, i.e., topologically structured sums of entities.

In the next section, we will propose an analysis that makes use of the mereotopological notions of MSSC and CLSTR in order to explain the meaning of collective and singulative morphology in Ukrainian.

4 Analysis

Drawing on the theory of Grimm (2012), we propose a mereotopological analysis on which the Ukrainian collective suffix -j- is modelled as a clustering device, whereas the Ukrainian singulative suffix -yn- is captured as a declustering atomizer (see also Scontras 2014). We postulate that both morphemes in question denote predicate modifiers that select for a specific type of input and apply the clustering or declustering operation, respectively. In particular, -j- takes a predicate of discrete
singu lar integrated wholes and turns it into an aggregate predicate, i.e., an expression denoting a property of individuals prototypically conceptualized as clusters. On the other hand, -yn- selects for an aggregate predicate and yields a predicate of integrated objects that originated as part of a cluster. We argue that our proposal provides an explanation of the semantic and distributional properties of the suffixes -j- and -yn- by utilizing the ontological distinction between integrated objects, arbitrary sums thereof and topologically structured clusters, as discussed in §3. Our aim is, thus, to demonstrate based on the Ukrainian data introduced in §2 how these conceptual notions are encoded in natural language semantics and impact grammar.

4.1 Three types of nominal predicates

Let us start by making a semantic distinction between three types of nominal predicates corresponding to singular count nominals, e.g., a pear, regular plurals, e.g., pears, and aggregate nouns, e.g., rice. First, let us consider concrete singular count nouns such as pear. We follow Grimm (2012) and Wągiel (2021b) in assuming that such expressions designate properties of MSSC individuals. Therefore, on our account the Ukrainian singular count noun hruška ‘pear’ has the semantics in (33). We model it as a predicate that is true of objects that are pears conceptualized as integrated wholes (represented as MSSC entities). For convenience, we will call such nouns OBJECT PREDICATES and use the abbreviation OBJ to refer to the corresponding higher-order property.

\[
\text{hruška} = \lambda x [\text{MSSC}(\text{PEAR})(x)]
\]

(33)

The second type of nominal expressions to be considered here are PLURAL PREDICATES (to be abbreviated as PL). These involve regular plurals of concrete count nouns such as pears. We assume that pluralization is achieved via the standard *-operator (Link 1983), which attaches to an object predicate and transforms the original set into a set containing the relevant MSSC objects as well as all the pluralities obtained by summing up these MSSC objects. For example, the Ukrainian plural noun hrušky ‘pears’ with the semantics in (34) denotes a set of pears conceptualized as integrated wholes along with all their sums.

\[
\text{hrušky} = \lambda x [\text{*MSSC}(\text{PEAR})(x)]
\]

(34)  

Finally, the third type of nominals to be discussed in this paper are uncountable granular nouns such as rice. These expressions denote properties of entities prototypically conceptualized as forming aggregates of objects, and thus we will call them AGGREGATE PREDICATES (with the abbreviation AGGR referring to the corresponding higher-order property). Building on the treatment of Slavic derived spatial collective nouns by Wągiel (2021a), we postulate that Ukrainian aggregate nouns
such as *rys ‘rice’* denote properties of clusters with the semantics such as the one in (35). The proposed denotation is intended to account for the intuition that aggregate nouns designate entities conceptualized as spatially structured configurations of individuals that typically either touch each other or remain in close proximity and often occur in large quantities. As such they are perceived in a different way than integrated objects, arbitrary sums thereof and homogeneous substances like liquids, gases and solid materials.

(35) \[ \text{rys} = \lambda x [\text{CLSTR}(\text{RICE})(x)] \]

On a conceptual level, the distinction between object, plural and aggregate predicates can be viewed as a linguistic reflex of different ways in which human beings perceive and conceptualize different types of part-whole structures. We argue that the contrast in the perception of integrated objects, arbitrary sums thereof and topologically structured clusters relates to the ontological distinction between the three types of entities introduced in §3. On our account, this intuition is captured by encoding a different part-whole structure in the meaning of object, plural and aggregate predicates. As a result, we will be able to capture the relevant semantic and distributional properties of the Ukrainian suffixes *-j-* and *-yn-* discussed in §2.

### 4.2 Deriving collectives

First, let us propose the semantics for derived spatial collectives. Drawing on Wagiel (2021a), we postulate that the Ukrainian collective suffix *-j-* is a clustering device.

We model it as a predicate modifier (type \(\langle\langle e, t\rangle, \langle e, t\rangle\rangle\)) with the semantics in (36). It takes an object predicate, i.e., an expression denoting properties of MSSC objects, and yields an aggregate predicate, which denotes a property of clusters. We capture its selectional restriction in terms of a definedness condition. In (36), *OBJ*\((P)\) is a presupposition that ensures that the input is an object predicate. The resulting expression denotes a set of clusters consisting of integrated objects in the relevant spatial configuration.

(36) \[ [-j-] = \lambda P: \text{OBJ}(P) \lambda x [\text{CLSTR}(P)(x)] \]

With the semantics in (36), let us demonstrate how our approach derives the meaning of spatial collectives formed with *-j-*. To illustrate, let us consider the collective noun *pirja ‘(clustered) feathers (as a mass)’*, introduced in (7a) and repeated here as (37).

(37) \[
\begin{align*}
\text{pero} & \quad \Rightarrow \text{pir-j-a} \\
\text{feather.SG} & \quad \text{feather-COLL-SG} \\
\text{‘a feather’} & \quad \Rightarrow \text{‘(clustered) feathers (as a mass)’} \quad \text{(Ukrainian)}
\end{align*}
\]
Clustering and declustering things

Since the base *pero* ‘a feather’ is a typical count noun, we assume it denotes the object predicate in (38). In other words, it is true of an entity that is a feather conceptualized as a discrete integrated whole.

(38) \[
\text{[pero]} = \lambda x [\text{MSSC(FEATHER)}(x)]
\]

When the suffix *-j-* attaches to *pero* ‘a feather’, the presupposition in (36) is satisfied since the base denotes an object predicate, recall the semantics in (36). The result of the semantic composition are the truth conditions in (39), specifically the collective *pirja* denotes a set of clustered entities that consist of feathers.

(39) \[
\text{[pirja]} = \text{[[-j-](pero)]} = \lambda P: \text{AGGR}(P) \lambda x [\text{CLSTR}(P)(x)](\text{[pero]}) = \lambda x [\text{CLSTR}(\text{[pero]})(x)]
\]

Our analysis explains the restricted distribution of the suffix *-j-*, namely the fact that it attaches only to countable nouns that typically denote relatively small entities that usually appear in large numbers and form clumps, piles, swarms and other types of topologically structured aggregates. The semantic derivation in (39) demonstrates how *-j-* clusters such entities in order to yield an aggregate predicate.

4.3 Deriving singulatives

Let us now move on to the semantics for Ukrainian singulatives. Following Grimm’s (2012) approach, we propose that the singulative suffix *-yn-* is a declustering device (similar to atomizers in Scontras’s 2014 account) with the meaning in (40). In particular, it denotes a predicate modifier that selects for an aggregate predicate, i.e., an expression denoting a property of clusters, and returns an object predicate, which denotes a property of MSSC objects. Again, we model the selectional restriction in terms of a presupposition. In (40), \( \text{AGGR}(P) \) is a definedness condition that guarantees the input to be an aggregate predicate. The output, on the other hand, is a set of singular integrated objects that originated as parts of a cluster. The resulting predicate can, thus, be pluralized and counted.

(40) \[
\text{[-yn-]} = \lambda P: \text{AGGR}(P) \lambda x \exists y [P(y) \land x \subseteq y \land \text{MSSC}(P)(x)]
\]

To illustrate how our approach captures singulative morphology, let us consider the semantic derivation of the singulative noun *popelyna* ‘a speck of ash’ introduced in (13a) and repeated here as (41).

(41) popil \( \Rightarrow \) popel-yn-a
ash.SG ash-SGV-SG
‘ash’ \( \Rightarrow \) ‘a speck of ash’ (Ukrainian)
For the uncountable base popil ‘ash’, we assume the aggregate semantics in (42). We argue that this is a linguistic reflex of the fact that ash, similar to, e.g., sand and rice, is perceived and conceptualized as an aggregate of externally connected particles rather than a homogeneous body of a substance.

(42) \([\text{popil}] = \lambda x [\text{CLSTR}(\text{ASH})(x)]\) \text{AGGR}

When the suffix -yn- combines with popil ‘ash’, the selectional restriction in (40) is met since the base is an aggregate predicate denoting a property of clusters. As a result, we get the denotation in (43). The singulative popelyna denotes a set of entities that are ash, but are at the same time conceptualized as MSSC objects, i.e., a set of separate discrete specks of ash.

(43) \([\text{popelyna}] = [\text{-yn-}][[\text{popil}]] = \\
= \lambda P: \text{AGGR}(P) \lambda x \exists y [P(y) \land x \sqsubseteq y \land \text{MSSC}(P)(x)][[\text{popil}]] = \\
= \lambda x \exists y [[\text{popil}](y) \land x \sqsubseteq y \land \text{MSSC}([[\text{popil}]])(x)]\text{OBJ}

The proposed analysis captures the effect of the singulative designating a unit within an aggregate. Moreover, it explains the constrained distribution of the suffix -yn-, i.e., the fact that it only combines with nouns designating clustered entities. In addition, the predicate resulting from declustering in (43) is countable and can be pluralized by means of the *-operator, recall (34), which would yield a set of individual MSSC specks of ash along with all pluralities obtained by sum formation.

4.4 Recursive shifting

An important advantage of the proposed system is that it allows for the alternating application of the clustering and declustering operations in any order. At the same time, it captures native speakers’ intuitions concerning subtle semantic differences between secondary singulatives and corresponding unit nouns, on the one hand, and secondary collectives and corresponding aggregate nouns, on the other. Let us first demonstrate how our approach accounts for the meaning of secondary singulatives. To illustrate, let us consider (20), repeated here as (44).

(44) pero ⇒ pir-j-a ⇒ pir-j-in-a
feather.SG feather-COLL-SG feather-COLL-SGV-SG
‘a feather’ ⇒ ‘(clustered) feathers’ ⇒ ‘a (single) feather’ (Ukrainian)

The process of deriving the secondary singulative in (44) starts with the object predicate pero ‘a feather’, recall (38), which serves as the input for the collective suffix -j- and the outcome is the aggregate predicate true of clustered feathers, recall (39). Finally, the singulative suffix -yn- attaches and as a result we obtain the
Clustering and declustering things

semantics in (45). The outcome is again an object predicate but unlike _pero_, the trace of clustering is preserved in its semantics, which accounts for native speaker intuitions regarding the subtle difference in meaning between the two.

\[
\begin{align*}
\text{[pirjina]} &= [-\text{yn-}][\text{[pirja]}] = [-\text{yn-}][-\text{j-}][\text{[pero]}] = \\
\lambda P: \text{AGGR}(P) \lambda x \exists y (P(y) \land x \subseteq y \land \text{MSSC}(P)(x))[[\text{pirja}]] = \\
\lambda x \exists y[[\text{pirja}](y) \land x \subseteq y \land \text{MSSC}[[\text{pirja}])(x)]
\end{align*}
\]

Let us now turn to secondary collectives such as (24), repeated here as (46).

\[
\begin{align*}
\text{popil} &\Rightarrow \text{popel-yn-a} \Rightarrow \text{popel-yn-nj-a} \\
\text{ash-SG} &\quad \text{ash-SGV-SG} \quad \text{ash-SGV-COLL-SG} \\
\text{‘ash’} &\Rightarrow ‘\text{a speck of ash’} \Rightarrow ‘(clustered specks of) ash’ \quad \text{(Ukrainian)}
\end{align*}
\]

In this case, the derivation begins with the aggregate predicate _popil_ ‘ash’, recall (42), to which the singulative suffix _-yn_- attaches in order to form the object predicate true of declustered specks of ash, recall (43). This in turn serves as the input for the suffix _-j-_, which triggers clustering and derives the aggregate predicate in (47).

\[
\begin{align*}
\text{[popelynnja]} &= [-\text{j-}][\text{[popelyna]}] = [-\text{j-}][-\text{yn-}][\text{[popil]}] = \\
\lambda P: \text{OBJ}(P) \lambda x [\text{CLSTR}(P)(x)][\text{[popelyna]}] = \\
\lambda x [\text{CLSTR}[[\text{popelyna}](x)]
\end{align*}
\]

A consequence of our approach is a semantics of nouns that is richer than typically assumed. In addition to singular objects and arbitrary sums thereof one needs to assume a category of clusters, i.e., structured configurations of plural entities. We believe that the data discussed in this paper further support proposals in the spirit of Grimm (2012) that postulate a mereotopological approach to nominal semantics.

5 Conclusion

In this paper, we investigated semantic properties of derived spatial collectives and singulatives in Ukrainian. We explored distributional constraints and semantic contribution of the collective suffix _-j-_ and the singulative suffix _-yn_- and the interaction between the two when they are stacked in secondary singulatives (_-j-_ + _-yn_-) and secondary collectives (_-yn_- + _-j_-). In order to account for the data, we proposed a mereotopological analysis that is based on the distinction between object predicates, i.e., expressions denoting properties of entities conceptualized as integrated objects, and aggregate predicates, i.e., expressions denoting properties of clusters, which are structured pluralities of entities that form spatial configurations of connected objects. We argued that the suffix _-j-_ introduces clustering by turning an object predicate into an aggregate predicates, whereas the suffix _-yn_- selects for an aggregate predicate and yields an object predicate, thus triggering individuation via declustering.
References


Clustering and declustering things


Wągiel, Shlikhutka


Marcin Wągiel
Department of Czech Language
Arna Nováka 1
602 00 Brno
Czech Republic
Centre for Corpus and Experimental Research on Slavic Languages “Slavicus” (SEI)
Kuźnicza 22
50-138 Wrocław
Poland
marcin.wagiel@phil.muni.cz

Natalia Shlikhutka
Centre for Corpus and Experimental Research on Slavic Languages “Slavicus” (SEI)
Kuźnicza 22
50-138 Wrocław
Poland
natalia.shlikhutka@uwr.edu.pl